

# LA CODIFICACIÓN DE LA INFORMACIÓN EN EL CD-DA

Jaume Llardén Prieto

**E**l standard Compact Disc se ha introducido masivamente y popularizado en el mercado en esta última década. Y no sólo por su uso como soporte musical, sino también por los diferentes formatos dirigidos al mundo informático (CD-ROM, *Compact Disc Read Only Memory*), a aplicaciones interactivas, que incluyen sonido, texto e imagen (fija o vídeo), como el DVI (*Digital Video*

*Interactive*) y el CD-I (*Compact Disc Interactive*), o a vídeo, como el single CDV (*Compact Disc Video*), de cara dorada en vez de plateada.

En este artículo se verá cuál es el proceso de codificación del audio y cómo se graba la información en el disco siguiendo el formato CD-DA. En los otros formatos mencionados el proceso de codificación es diferente, si bien la grabación física en el disco de la información digital (el CDV contiene informa-

ción analógica y digital) es igual que en el CD-DA.

## 1. ¿Qué es el CD-DA? Standard CD-DA.

El Compact Disc-Digital Audio es un sistema de almacenamiento digital de información en disco óptico desarrollado por Philips y por Sony Corporation en 1980, capaz de almacenar 74 minutos y 33 segundos de música de alta fidelidad (hasta 80 minutos en algunas variaciones).

Philips y Sony, y en el standard BNN15-83-095 "Compact Disc Digital Audio System" del IEC (*International Electrotechnical Commission*), y se dan en la tabla de la figura 1.

## 2. ¿Cómo se lee y se graba la información digital?

El sistema CD-DA está diseñado para guardar la cantidad ingente de información que produce más de una hora de audio con calidad hi-fi.

El disco CD-DA tiene 4 capas de material, que se nombrarán desde la cara de lectura a la cara de la etiqueta. La mayor parte de los 1.2 mm de grosor del disco está formada, usualmente, por policarbonato inyectado, aunque pueden usarse otros

- |  |
|--|
| <ul style="list-style-type: none"><li>· <i>Características generales del disco</i><ul style="list-style-type: none"><li>· Tiempo máximo de reproducción : 74 minutos 33 segundos.</li><li>· Sentido de la rotación : el de las agujas del reloj mirando la cara de la etiqueta.</li><li>· Velocidad lineal de rotación : un valor entre 1.2 y 1.4 ms<sup>-1</sup>, mantenido constante durante la lectura.</li><li>· Velocidad radial de rotación : 500-200 rpm</li></ul></li><li>· <i>Características físicas del disco</i><ul style="list-style-type: none"><li>· Diámetro, peso y grosor del disco : 12 cm, de 14 a 33 g y 1.2 mm</li><li>· Diámetro del agujero central : 15 mm</li></ul></li><li>· <i>Características del sistema óptico</i><ul style="list-style-type: none"><li>· Longitud de onda standard en el aire : <math>\lambda_0 = 780</math> nm</li><li>· Anchura focal : <math>\pm 2</math> <math>\mu</math>m</li></ul></li><li>· <i>Características del formato de la señal</i><ul style="list-style-type: none"><li>· Número máximo de tracks (canciones) : 99 tracks.</li><li>· Cuantificación y frecuencia de muestreo : cuantificación lineal de 16 bits a 44.1 kHz</li><li>· Velocidad de transferencia del disco al lector : 4.3218 Mbits/s</li><li>· Velocidad de transferencia de datos : 1.4112 Mbits/s o 172.26 kbytes/s</li><li>· Código de corrección de errores : CIRC (<i>cross interleave Reed-Solomon code</i>).</li><li>· Sistema de modulación : EFM (<i>eighth-to-fourteen modulation</i>).</li></ul></li></ul> |
|--|

Figura 1.- Características del standard CD-DA.

Es un sistema digital porque la información a grabar requiere un muestreo y una cuantificación, y el medio es óptico porque la información grabada en el disco se lee con un haz de luz, no como en los diskettes de ordenador (medio magnético) o los antiguos LP (medio mecánico).

Los parámetros del standard están definidos en el Libro Rojo de

materiales plásticos transparentes (ver p. 72, [1]). Esta capa se denomina *sustrato*. Sobre el sustrato hay una *superficie reflectante metalizada* (de entre 50 y 100 nm de grosor), que suele ser de aluminio, donde se hacen unos bultos o agujeros (en inglés *bumps* y *pits*), de longitud variable. (Llamarlos bultos o agujeros depende de la perspectiva, si se mira desde la

JAUME LLARDÉN PRIETO es estudiante de quinto curso de Ingeniería Técnica Superior de Telecomunicaciones de Barcelona (UPC).

cara de lectura o la cara de la etiqueta). La distancia entre bultos consecutivos, llamada tierra (en inglés *land*), es también variable. Los bultos y las tierras codifican la información, como se verá. Sobre la superficie metalizada hay una *capa superior* protectora (la capa protectora tiene entre 10 a 30  $\mu\text{m}$ ). Finalmente, encima, se encuentra la *tinta de la etiqueta* del disco (de unos 5  $\mu\text{m}$  de grosor).

¿Cómo se lee el CD-DA? La lectura se basa en la interferencia de la luz. El rayo láser se enfoca sobre el camino de bultos impreso en el disco, y al ir iluminando bultos y tierras la intensidad del rayo reflejado varía (se modula). Hay que aclarar que el láser, cuando ilumina el bulto, *también* ilumina una zona adyacente de tierra, hecho imprescindible para lograr la lectura, como se comentará inmediatamente (ver figura 2).

Para comenzar, supongamos que los bultos tienen exactamente una altura de  $\frac{1}{4}$  de la longitud de onda en el policarbonato (lo que no es del todo cierto). La luz que se refleja en la tierra viaja  $\frac{1}{2}$  ( $\frac{1}{4} + \frac{1}{4}$ ) de longitud de onda más que la luz que se refleja en el bulto. Esto crea una diferencia de fase entre la luz proveniente del bulto y su tierra adyacente, produciendo una interferencia destructiva que hace que se cancelen. Teóricamente la cancelación es total. En la práctica, el haz de laser es excesivamente ancho para que se produzca esta cancelación, y los bultos se hacen algo menores de  $\frac{1}{4}$  de longitud de onda. De esta manera los datos codificados en el disco pueden ser recuperados y convertidos a una señal eléctrica. Ha de notarse que la señal de audio no está grabada en la superficie del disco, como en los LP, sino protegida dentro del disco, y leída por un haz de luz que desgasta

el disco tanto como se desgastan las palabras de un libro al ser leídas.

¿Y cómo se guarda la información? Contrariamente a lo que se podría esperar, los bultos y tierras no son directamente ceros y unos. Son los bordes de los bultos lo que nos indican los unos, y la longitud del área entre los bordes lo que codifica los ceros. Esto se puede ver en la figura 3. También está indicado en la figura el tiempo  $T_b$  que se tarda en

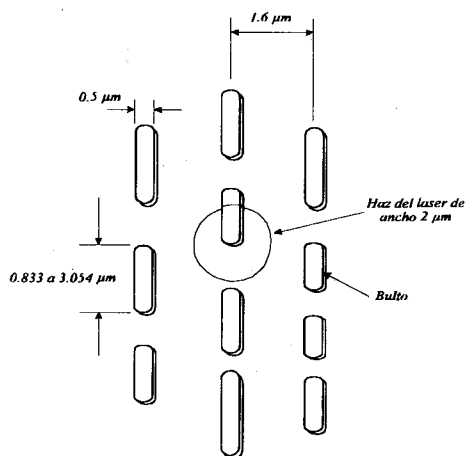


Figura 2.- Superficie metalizada reflectante vista desde la cara de lectura.

leer los bits, que vale  $T_b = 1/(\text{velocidad de transferencia del disco al lector}) = 1/4321800 = 231.38 \text{ ns}$ .

### 3. ¿Cómo se codifica la información digital?

El standard CD-DA muestrea el canal derecho e izquierdo a 44.1 kHz y cada muestra, derecha e izquierda, es codificada con 16 bits. Las muestras se ordenan anteponiendo las del canal izquierdo a las del derecho.

Obviamente, no se pueden grabar estos bits directamente sobre el disco. Se necesitan técnicas para controlar los errores (en inglés *dropouts*) de lectura producidos, generalmente,

por la impresión incorrecta de los datos, arañazos en el disco o problemas de pérdida de sincronismo en el lector. Los dos tipos de errores que aparecen en el CD-DA son:

*errores aleatorios*, definidos como errores que se presentan aislados o en grupos de no más de  $17T_b$ . Un caso típico es un bulto mal impreso en el disco.

*errores de ráfaga* que se presentan en grupos mayores de  $17T_b$ . Pueden ser producidos, por ejemplo, por un arañazo.

En el CD-DA se usan bits de paridad (los llamados bytes P y Q) y entrelazado de la información, que siguen el código CIRC, y modulación EFM.

¿Para qué se usan los bytes de paridad P y Q? Se calculan a partir del código Reed-Solomon (ver p. 92, [1] y p. 63 [2]) e introducen redundancia para la corrección de errores aleatorios.

¿Qué se persigue con el entrelazado? El entrelazado es un reordenamiento de la información y se realiza con la idea de que, al reponer las muestras en sus posiciones originales, los errores del disco quedarán distribuidos entre datos válidos, siendo muy probable poder corregirlos con los bytes de la paridad.

¿Qué se intenta con la modulación? Pasar a combinaciones de bits con pocas transiciones cero-uno para facilitar su grabación física en el disco.

¿Cómo se llega al formato final de la información a partir de las muestras de ambos canales? La unidad básica de información son 12 muestras de sonido (cada una de 16 bits, como se ha visto), 6 muestras del canal derecho y 6 del canal izquierdo, que son  $6 \times 2 \times 16 = 192 \text{ bits} = 24 \text{ bytes}$ . Estos 24 bytes forman un bloque, que

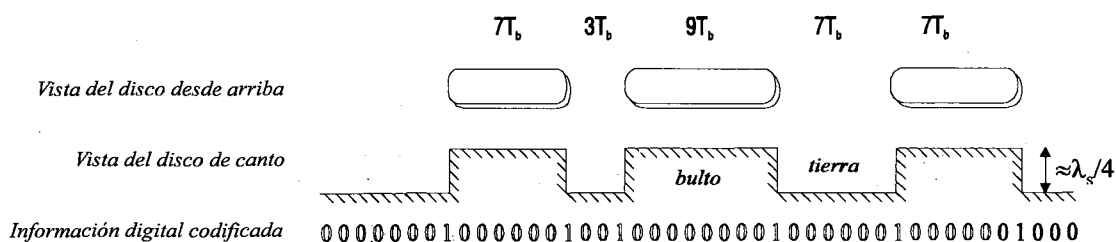


Figura 3.- Grabación de los ceros y unos a partir de los bultos y tierras ( $\lambda_s$  es la longitud de onda del sustrato).

llamaremos *inicial*, y se codifican según CIRC.

### 3.1 Codificación CIRC.

El Cross Interleave Reed-Solomon Code (CIRC) es un método de detección y corrección de errores que básicamente consiste en añadir 8 bytes de paridad, según la codificación Reed-Solomon, en dos pasos, llamados C1 y C2 (4 bytes en cada paso), separados por un entrelazado. El cálculo de la paridad es idéntico en ambos pasos. La diferencia reside en cuántos bytes se usan para obtenerlos. El C2 usa 24 bytes (los 24 del bloque inicial) y la C1 usa 28 bytes (24 del bloque inicial más los 4 bytes de paridad C2). El CIRC consigue una probabilidad de error de  $10^{-11}$ .

La codificación Reed-Solomon permite corregir hasta 2 bytes de una palabra de 28 o 32 bytes. ¿Cómo se logra?

Primero recordemos el concepto de distancia mínima  $d_{min}$  de un código: es el número *mínimo* de símbolos que se han de cambiar en una palabra de código para pasar de una palabra válida a otra también válida (no necesariamente contiguas). Un ejemplo de código con palabras de 6 símbolos (en este caso se tiene que cada símbolo es un bit) es:

|                             |            |
|-----------------------------|------------|
|                             | X1=101 000 |
| combinaciones<br>no legales | Y1=101 001 |
|                             | Y2=101 010 |
|                             | Y3=101 100 |
|                             | Z1=101 101 |
|                             | Z2=101 110 |
|                             | Z3=101 011 |
|                             | X2=101 111 |

Las palabras del código son  $X_1$  y  $X_2$ , y como se observa  $d_{min}=3$ . Si recibimos cualquiera de los códigos  $Y_i$  (ilegales) se considera un  $X_1$  incorrecto, ya que sólo varían un símbolo (1 bit) respecto  $X_1$ , pero dos símbolos respecto  $X_2$ . Es decir, que con  $d_{min}=3$  sólo podemos reconocer un símbolo erróneo por palabra, es decir, podemos corregir el error de un símbolo en una palabra. La relación general, si  $n_e$  indica el número de símbolos erróneos que se quieren corregir en una palabra, es  $d_{min}=2n_e+1=n_e+(n_e+1)$ . Esta expresión se deduce de que la palabra leída incorrectamente ( $Y_i$ ) ha

de estar  $n_e$  símbolos de distancia de la correcta ( $X_1$ ) y a una distancia mayor que  $n_e$  de otra palabra legal del código ( $X_2$ ), como mínimo a  $n_e+1$ .

¿Y cómo se pueden corregir dos símbolos erróneos? Con  $d_{min}=5$ , que es precisamente la que tiene la codificación Reed-Solomon. En este caso los símbolos son bytes y la palabra, a la que añadir la paridad, es de 24 (C2) o 28 (C1) bytes. Cada palabra (sin paridad) legal varía de una a otra en por lo menos un símbolo (1 byte). Así que  $d_{min}=1$ . Es la paridad la que da la distancia deseada: el código Reed-Solomon calcula los 4 bytes de manera que varíen de una palabra a otra y permitan  $d_{min}=5$ . Esto permite finalmente corregir 2 bytes en palabras de 28 y 32 bytes, que son las de 24 y 28 bytes más los 4 bytes de paridad.

Hay otro procedimiento importante que permite corregir hasta 4 bytes (en general  $d_{min}-1$ ) en una palabra de 28 o 32 bytes. La condición para aplicar este procedimiento es restrictiva: hay que conocer a priori en qué posiciones de la palabra están los bytes erróneos. Supongamos que nos llega una palabra  $Y_1$  con 4 bytes marcados como erróneos. Dado que la distancia entre palabras de código válidas  $X_i$  es cinco, sólo habrá una palabra  $X_1$  en la que coincidan los bytes válidos, y  $Y_1$  se considerará un  $X_1$  erróneo. Las posiciones con bytes erróneos se llaman posiciones de borrado (*erasure position*). Más adelante veremos que si es posible conocer las posiciones de bytes incorrectos y que este método es muy útil.

La secuencia completa de la creación de una trama CIRC es (ver figura 4): a) retardo y ordenación, b) codificación Reed-Solomon C2 (cálculo de la paridad Q), c) entrelazado, d) codificación Reed-Solomon C1 (cálculo de la paridad P) y e) retardo con inversión lógica de paridad. En la decodificación la secuencia es la inversa. En el CIRC los bloques de 24

bytes se codifican en paralelo, es decir, se tratan los bytes de 24 en 24.

#### Retardo y ordenación.

Primero se retardan las muestras pares dos bloques. Cuando se habla de retardar muestras en  $n$  bloques se ha de entender que es un retardo de la señal digital de  $n \times D$ , con  $D = 1 / [88200 \text{ (muestras / segundo)}] \times 12 \text{ (muestras / bloque)} / [24 \text{ (bytes / bloque)}] = 5.67 \mu\text{s}$ . Más adelante los bloques tendrán 28 y 32 bytes, números que han de sustituir a 24 en el anterior retardo.

El siguiente paso es un reordenamiento (simétrico entre las 6 primeras y las 6

últimas muestras) para separar las pares de las impares. Esto permite, en la decodificación, distribuir entre bytes correctos los bytes erróneos no corregibles con las paridades P y Q (pero marcados como erróneos por ellas) para su interpolación.

#### Codificación C2 (adición de la paridad Q).

Ahora los 24 bytes del bloque ordenado se usan para calcular 4 bytes de paridad, llamada paridad Q. La paridad Q está diseñada para corregir los errores de ráfaga marcados por la paridad P y también los errores aleatorios que no se pudieron corregir con ella. Los 4 bytes Q se sitúan en el centro de los 24 bytes del bloque primario para aumentar la distancia entre muestras pares e impares, y así mejorar la interpolación en caso de una ráfaga de errores. A este bloque de 28 bytes le llamaremos bloque codificado C2.

#### Entrelazado.

Después se hace un entrelazado cruzado: los 28 bytes que forman el bloque codificado C2 se retardan con diferentes períodos múltiplos de 4 bloques: el primer byte del bloque no se retarda, el segundo se retarda 4 bloques, el tercero 8 bloques y así hasta el byte vigésimo octavo, que se retarda en 108 bloques. De esta manera cada uno de los 28 bytes se almacena en otros tantos bloques dis-

*La codificación Reed-Solomon, usada en el CD-DA, permite corregir hasta 2 bytes de una palabra de 28 o 32 bytes.*

tribuidos entre 109 bloques y los 28 bytes del bloque resultante proceden de 28 bloques codificados C2 diferentes. Este entrelazado está diseñado para, en la decodificación, aislar los errores de ráfaga, es decir, convertirlos en errores aleatorios. Al decodificar primero se usa la paridad P, se desentrelaza, y luego se usa la Q. Por tanto es la paridad Q la que puede corregir los errores de ráfaga que en el paso anterior (el entrelazado) son convertidos en errores aleatorios.

#### Codificación C1 (adición de la paridad P).

De los 28 bytes resultantes se calculan 4 bytes más de paridad según CIRC C1. La paridad P corrige la mayor parte de errores aleatorios y detecta y aísla los errores de ráfaga para que los corrija la paridad Q. En el decodificador, si se detectan más errores, se marca todo el bloque con un *erasure flag* para que la paridad Q actúe sobre ellos.

#### Retardo con inversión lógica de paridad.

Los bytes impares se retardan en un bloque. Esto evita que los errores aleatorios corrompan más de un byte por bloque; más adelante (sección Bits de unión) comentaremos el porqué. Los bytes de paridad P y Q,

como último paso, son invertidos lógicamente para tener siempre bytes diferentes de cero. Esto ayuda en la lectura de datos cuando hay zonas de silencio de audio. El bloque de 32 bytes resultante es la *trama CIRC*.

Para entender cómo funciona la corrección de errores en CIRC vamos a estudiar dos casos de lectura incorrecta de un bloque del disco: que se lean dos bytes contiguos erróneos (error aleatorio) y que se lean 23 bytes contiguos erróneos (error de ráfaga).

**Dos bytes.** Después de leer la información del disco, de desechar los bits de unión y de demodular la EFM (más adelante se explicarán estos pasos), el bloque de 32 bytes se retarda y se invierten las paridades. El retardo hace que los bytes dejen de ser contiguos y haya solo un error por bloque. El bloque pasa al decodificador C1. Al comprobar los bytes de paridad P detecta un byte erróneo. El byte podría ser un byte de paridad, sin merma de efectividad. Como hemos comentado antes, el código Reed-Solomon permite corregir hasta 2 bytes, así que corrige el byte detectado. El bloque se desentrelaza. En el decodificador C2

se quita la paridad Q y en el último paso se reordena y retarda, para obtener 24 bytes de datos correctos.

**23 bytes.** En este caso, después del retardo tenemos un bloque con 11 bytes erróneos y otro con 12 bytes erróneos. El decodificador C1 detecta más de 2 bytes erróneos por bloque y marca cada bloque con un *erasure flag*, sin poder corregirlos. Los bytes erróneos, por tanto, están indicados como tales para el resto de etapas. El desentrelazado es lo que ayudará a arreglar este panorama. Al desentrelazar los retardos que se aplican son los inversos que en el entrelazado, es decir, el primer byte se retarda 108 bloques, el segundo 104, y así hasta el último, que no se retarda. El resultado es que los bytes erróneos (marcados con un *erasure flag* y por tanto en posiciones conocidas) quedan distribuidos entre bytes correctos, y los bloques resultantes tienen menos de 3 bytes erróneos, y son corregibles por el decodificador C2. Una pregunta natural surge: ¿hasta cuántos bytes erróneos contiguos se pueden corregir con CIRC? Hasta 16 bloques o 512 bytes. Cuando el error es mayor se opta por la interpolación de los bytes erróneos si éstos, que están marcados y por tanto en posiciones conocidas, se encuen-

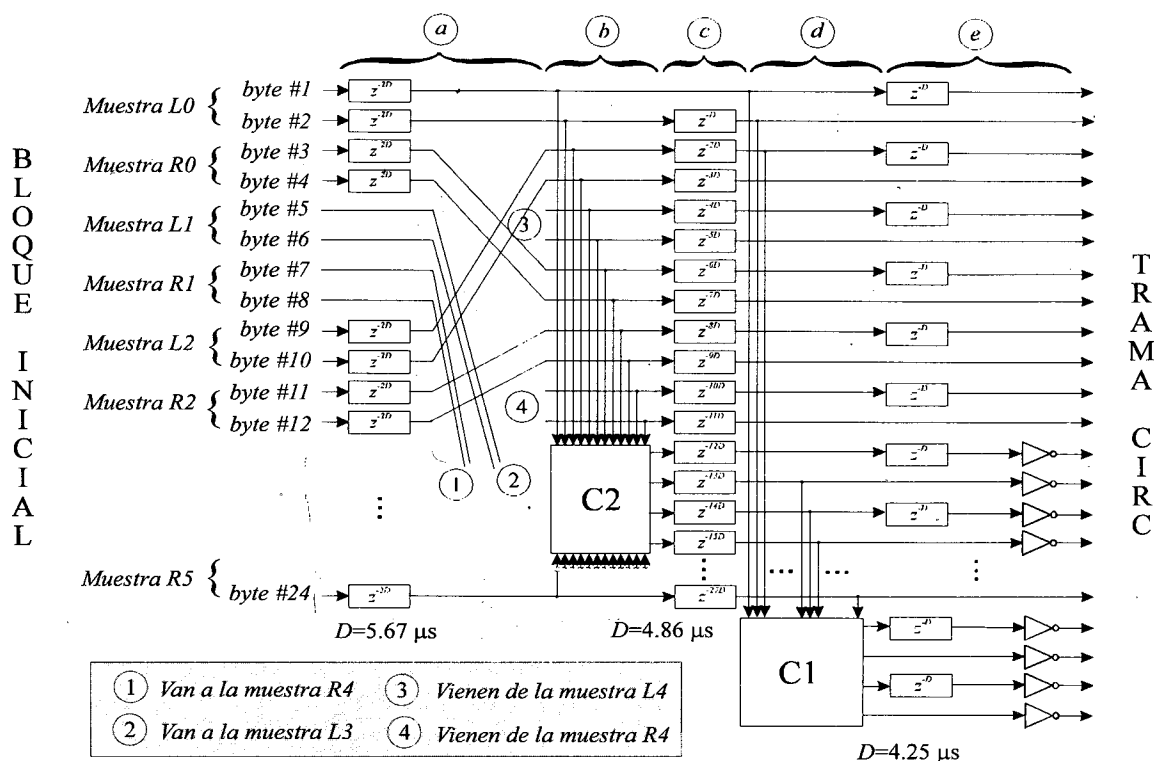


Figura 4.- Creación de la trama CIRC a partir del bloque inicial (en cada etapa el retardo  $D$  tiene un valor diferente).

tran entre datos correctos. Si incluso la interpolación es imposible, el decodificador opta por pasar los bytes erróneos como silencio. Estos silencios suelen ser muy cortos y normalmente no se notan.

### 3.2 Subcódigo.

Después de la codificación CIRC se añade un byte de subcódigo por bloque de 32 bytes. Los 8 bits de subcódigo se llaman P, Q, R, S, T, U, V y W. El CD-DA sólo usa los bits P y Q, que no tienen nada que ver con las paridades P y Q. Estos bits de subcódigo incluyen información sobre el número de canciones del disco, su comienzo, final y duración, el *copyright*, etc. Los bits de subcódigo se usan agrupando 98 bloques para obtener  $98 \times 8 = 784$  bits. Una explicación de cómo se usan se puede encontrar en [1], pág. 113.

### 3.3 Modulación EFM.

EFM quiere decir *modulación de ocho a catorce*: los 33 bytes (264 bits) escogidos en grupos de 8 en 8 bits se traducen a grupos de catorce bits, llamados *channel bits*. ¿Qué factores se tienen en cuenta para escoger las palabras de 14 bits? La asignación debe ser biyectiva y debe minimizar el número de transiciones entre ceros y unos.

Para que los bultos y tierras tengan longitudes controladas, se exige que las palabras de 14 bits tengan más de 2 pero menos de 10 ceros seguidos. De las posibles  $2^{14}=16384$  combinaciones 267 satisfacen este criterio. Sólo se necesitan  $2^8=256$ .

### 3.4 Bits de unión.

Los grupos de 14 bits se unen con 3 bits. Dos de los bits de unión son siempre 0, para evitar tener unos sucesivos entre las palabras de 14 bits. El

tercer bit, que puede ser cero o uno, se añade para ayudar a la sincronización del reloj y para reducir los componentes de baja frecuencia de la señal digital. En la demodulación los tres bits se unión se desechan.

Cada byte de la trama CIRC de 32 bytes, al modularlo y añadirle los bits de unión, pasa a tener 17 bits. Este grupo de 17 bits es el que se grabará en el disco. Recordemos que los errores aleatorios estaban definidos como aquellos que como máximo tenían 17 bits. Es obvio entonces que un error aleatorio como máximo afectará a dos símbolos contiguos de 17 bits. El codificador CIRC retardaba

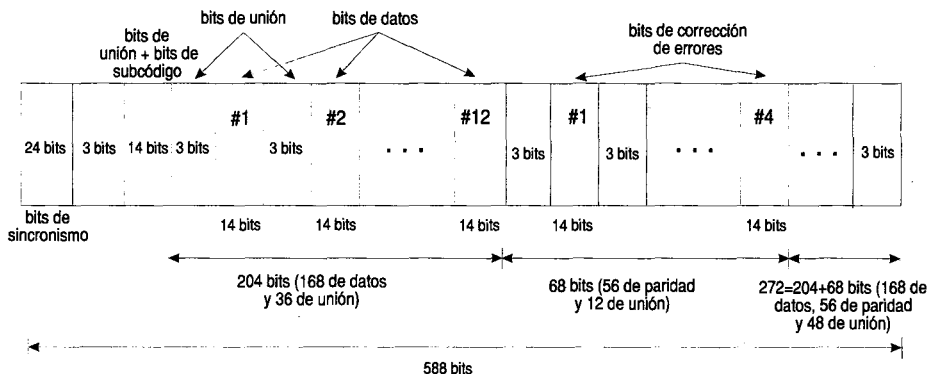


Figura 5.- Trama final de la información que se graba en el CD-DA.

los bytes impares en 1 bloque como último paso antes de invertir las paridades P y Q. Así que al leer el disco, si se comete un error de 17 bits, después de la demodulación EFM y de desear los bits de unión, los dos bytes contiguos se asignarán a bloques diferentes al ser retardados los símbolos pares. Por tanto como máximo, ante un error aleatorio, se pierde un byte por bloque. Si ocurre un error de ráfaga se perderán más bytes, y la paridad Q, en la medida de los posible, los corregirá.

### 3.5 Adición de la sincronización y formato de la trama.

Los grupos de 564 bits ( $33 \times 14$  channel bits +  $34 \times 3$  bits de unión) resultantes deben ser diferenciados o delimitados unos de otros. Se usa por tanto una palabra de 24 bits de sincronización al inicio de cada grupo, que es única y no puede confundirse con ningún dato. La palabra en cuestión, a la que se habrá de añadir los bits de unión, es 1000 0000 0001 0000

0000 0010: 3 transiciones separadas por 10 ceros. El formato de la trama final que se graba en el disco está formado por 588 bits: 24 bits de sincronismo, 336 ( $24 \times 14$ ) bits de datos, 112 ( $8 \times 14$ ) bits de corrección de errores, 14 bits de subcódigo y 102 ( $34 \times 3$ ) bits de unión. La organización de los bits se puede ver en la figura 5. Estos bits se imprimen en el policarbonato siguiendo la forma de la figura 3.

### 3.6 Velocidad de transferencia de la información.

El muestreo se hace a 44.1 kHz, obteniendo de cada muestra 32 bits, 16 del canal izquierdo y otros 16 del derecho. Por lo tanto, la fuente de datos a una velocidad de  $44.1 \times 10^3 \times 32 = 1411200$  bits/s o 172.26 kbytes/s (1 kbyte = 1024 bytes).

Esta será la velocidad de transferencia que tendrán los datos (la música) una vez leídos y decodificados en el lector de CD-DA.

La velocidad de transferencia de la trama final del disco al lector es mayor, ya que en el mismo tiempo tenemos que enviar mucha más información. La trama se crea en todo el proceso de codificación anteriormente explicado. En este proceso hemos pasado de los 192 ( $24 \times 8$ ) bits de datos iniciales a los 588 bits del formato final, de manera que la velocidad de transferencia de bits al lector hecha por el cabezal son  $1411200 \times 588 / 192 = 4321800$  bits/s, o 527.5 kbytes/s.

### 4. Bibliografía.

- [1] HEITARO NAKAJIMA Y HIROSHI OGAWA: *Compact Disc Technology*, Omsa Ltd., 1992
- [2] KEN C. POLHMANN: *The Compact Disc: a handbook of theory and use*, A-R Editions Inc., 1989